我国科研诚信政策变迁计量分析*

■ 冯凌子[1,2] 刘敬[3,4] 袁军鹏[1,2]

[1] 中国科学院文献情报中心 北京 100190 [2] 中国科学院大学经济与管理学院图书情报与档案管理系 北京 100190
[3] 中国科学院信息工程研究所 北京 100093 [4] 中国科学院大学网络空间安全学院 北京 100049

摘 要：[目的/意义]对我国中央发布的科研诚信政策和地方响应情况的现状及变迁作了分析，以求把握政策发展态势与阶段重点，为科研诚信建设提供支撑，为我国科研诚信政策发展研究提供帮助。[方法/过程]搜集整理108份与科研诚信相关的中央政策及80份参照落实中央政策的地方政策，使用政策文本计算与政策文献量化方法，对政策发布时间、发布机构、政策内容及政策响应情况的现状及其变迁进行分析梳理，得到了政策发布的时间特征、重要机构特征、内容侧重点变化以及政策响应的特征。[结果/结论]①科研诚信相关治理措施早在1980年就有涉及，近年来围绕国家创新建设的需求发文数量骤增，治理的态度更坚决，措施办法也更加具体。②科研诚信的治理由中央政府领导，教育部与科技部是各个阶段的主要管理部门，相关部门逐渐加入进行共同治理，目前已形成合力的格局。③各部门的科研诚信治理着力点、对象与方法略有差异，但总体而言，从方法与目标上来看，逐渐从自我约束、道德教育转变到制度建设与道德建设齐抓共管。④全国范围内绝大多数的地方省市都积极响应了中央政策，采取措施进行科研诚信治理，但相对而言有一定的时滞。

关键词：科研诚信 政策文本计算 政策文献量化 政策变迁 政策分析
分类号：G250
DOI：10.13266/j.issn.0252-3116.2020.09.009

1 引言

近年来，随着越来越多的科研不端事件被曝光，科研诚信这一议题也越来越受关注。有很多案例引起了科学界甚至整个社会的广泛热议。如2016年，学界与社会主要对"韩春雨事件"[1]中其研究是否属于科研不端展开了热烈的讨论；2017年，*Tumor Biology* 期刊因虚假同行评议一次性撤销了中国学者发表的107篇论文引起热议[2]；2018年，南京大学梁莹教授120多篇论文凭空消失，她的大量论文多涉嫌抄袭、一稿多投等科研不端问题，因而引起科学界和社会热议，并最终经过调查被记过、撤职[3]；2019年初，演员翟天临因一句"不知知网"引发一系列民间与官方调查，最终因为论文存在学术不端情况被撤销博士学位[4]。我国政府从20世纪90年代开始重视科学界的学术道德和学风建设，2005-2006年间先后出现的韩国"黄禹锡事件"和我国"汉芯"事件，促使有关部门更加重视并全面加强了科研诚信建设工作。2010年，时任国务院副总理的刘延东在科研诚信与学风建设座谈会上就提出"要强化监督管理，加大对科研不端行为的惩治力度，决不让弄虚作假、剽窃抄袭的行为有立足之地"[5]。2017年，习近平总书记在十九大报告中多次提到"创新"，强调要"推进诚信建设"[6]。2019年，李克强总理在第十三届全国人民代表大会第二次会议上的《政府工作报告》中也提到，要"惩戒学术不端，力戒浮躁之风"[7]。从政策层面来看，多年来，我国政府也发布了多项政策，对科研诚信问题进行治理。

政策文献是政府从事社会管理工作的基本反映[8]，传统的政策研究方法依靠专家的社会历史和知识背景，解读政策文本；新兴的政策研究方法，例如政策文献量化[8]、政策文本计算[9]等引入统计学、文献计量、内容分析、网络分析、文本挖掘等方法，降低了定性

* 本文系2018年"中国科学院文献情报领域引进优秀人才计划"择优支持项目（科发规函字〔2018〕12号）和科技部科技创新战略研究专项"科技伦理体系研究"（项目编号：ZLY201950）研究成果之一。
作者简介：冯凌子（ORCID：0000-0003-3410-9023），博士研究生；刘敬，硕士研究生；袁军鹏（ORCID：0000-0003-2803-5312），研究员，博士，博士生导师，通讯作者，E-mail：yuanjp@mail.las.ac.cn。
收稿日期：2019-05-15 修回日期：2019-10-27 本文起止页码：73-84 本文责任编辑：杜杏叶

的政策解读所造成的不确定性和主观性。对政策文献进行量化、计算可以了解政策的主体特征、政策现状与主题变迁等内容。目前研究的成果已较为丰硕，例如，黄萃等对新中国成立以来的科技政策作了政策计量分析[10]；张剑等以科技成果转化政策为例，对中国公共政策扩散的状况进行了文献量化研究[11]；李燕萍等基于政策文本，对改革开放以来我国科研经费管理政策的变迁、评介与走向作了内容分析[12]；杨凯瑞等基于国务院及 16 部委的数据，对中国中央政府支持创新创业发展政策文本进行了量化研究[13]；X. Zhang 等对我国农村信息化政策的演化作了政策计量分析[14]；黄萃等还从政策工具视角的对中国风能政策文本进行了量化研究[15]。但是目前对我国科研诚信政策进行量化分析的研究还比较少，目前也暂未发现有关地方对于中央科研诚信政策响应状况的研究。

目前我国对科研诚信已有的研究更多集中在相关概念、表现形式、影响因素、制度建设与评价体系上[16]，已有的一些基于政策文本计算或政策文献量化对科研诚信政策进行研究的文献中，有的对美国治理科研不端行为的政策进行量化分析，以探讨科研不端行为的治理政策[17]；有的对我国科研诚信政策的特征进行分析，对科研诚信的建设提出建议[18]。总体而言，已有研究中对我国科研诚信政策的变迁进行分析还比较缺乏。分析政策的变迁可以了解内化于政府文献中政策目标与政策主题的变化[19]，是了解政府执政理念转变脉络的方式。同时，对地方关于中央政策的响应情况进行分析，也有助于了解中央政策向地方推行扩散的状况。

本文使用政策文本计算与政策文献量化方法对我国中央发布的科研诚信政策和地方响应情况的现状及变迁作了分析，以求把握政策发展态势与阶段重点，为科研诚信建设提供支撑，为我国科研诚信政策发展研究提供帮助。

2 数据来源及分析方法

2.1 中央政策数据获取及分析

2.1.1 中央政策文本获取

近年来对于科研诚信相关问题我国出台了一系列的文件，包括法律法规、规章制度以及规范性文件等政策，这些文件中，有包含了科研诚信相关规定的纲要性文件，也有与科研诚信直接相关并提出办法措施与意见的文件。本文的政策文本数据有以下来源：①2017 年科学技术部科研诚信建设办公室编纂的《科研诚信建设相关法律法规和文件汇编》[20]整理了近年来 80 余份与科研诚信相关的文件，具有一定的权威性。本文以该书为基础，获取全书收录的政策文本作为分析数据，并在此基础上扩展搜集最新的相关文件。②从目前国内最为权威的法律法规信息检索系统“北大法宝”（http://www.pkulaw.cn）中获取政策文本，以与科研诚信主题最为紧密相关的“科研诚信”“科研不端”“学术不端”“学术道德”“学术规范”“学风建设”为关键词进行政策检索，下载政策全文，作为补充数据集。③为了使数据集的分析结果更准确，删去数据集中实质性内容较少的各机构内部工作会议的文本。最终得到 108 份文件，所有文本数据获取于 2019 年 2 月。

2.1.2 政策文本结构化信息组织

（1）文件题录信息获取。本文从北大法宝数据库中获取中央政策的结构化题录信息，该题录信息包括标题、时效性、效力级别、发布部门、发文字号、发布日期、实施日期等数据，对于数据库中未收录的政策，题录信息由人工按照北大法宝数据库中的信息结构完善补齐。

（2）中央部委发文机构规范。对于发文机构名称的规范作以下处理办法：①由二级机构（如“办公厅”“组织部”）发布的文件，统一标注一级机构名称，例如“中共中央办公厅”与“中共中央组织部”将机构名标注为“中共中央”。②国务院机构中，国务院办公厅机构名称规范为“国务院”，涉及组成部委的文件，保留到部委一级名称，例如“科技部科研诚信办公室”规范为“科技部”。③国务院直属的事业单位、部委管理的国家局保留到具体单位一级的名称，如由国家卫生健康委员会管理的国家中医药管理局，依然保留国家中医药管理局一级名称。④已撤销或变更的部委机构发布的文件，统一规范为十三届全国人大一次会议全体会上国务院机构改革方案通过后的机构名称，原发文单位名称的具体规范名称对应关系，经查找现行政府机构的官方网站中“机构沿革”确定。

（3）文件发布时间确定。文件的时间包含发布时间、实施时间，以及部分文件的修订时间与实施时间。本文中将没有修改版本的文件发布时间上缩至“年”，为文件发布的当年。经过修订的法律法规，则根据修订的内容是否与科研诚信直接相关，来决定是否同时保留最初发布年份与修订年份，如修订内容直接相关，则同时保留最初发布年与修订年，如不直接相关，则仅保留最初发布年份。

（4）失效文件处理办法。由于部分搜集到的政策

文件已失效，为了体现各年法律法规、政策发布的动态，以及分析政策内容的变迁，因此将已失效的文本也列入了分析数据集中。

2.1.3　政策文本内容变迁分析数据及方法

为了发现我国科研诚信政策的主要内容，以及各阶段、重要机构内容变迁，本文对高频关键词进行聚类，并分析对比了样本总体及部分重要中央机构发布的政策中的关键词，且分时间段对比了关键词的变化。方法步骤如图1所示：

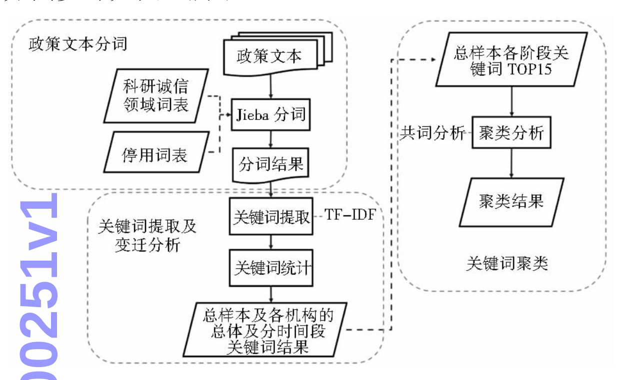


图1　政策文本内容变迁分析数据及方法

注：总样本即所有中央科研诚信政策文本

（1）政策文本分词。研究中使用了 jieba 分词工具（https://github.com/fxsjy/jieba）进行中文分词，pandas 数据统计工具包（http://pandas.pydata.org）进行统计，团队对大量科研诚信领域的论文进行了研读，根据研读成果整理出了自建词表“科研诚信领域词表”，并将其用于政策文本分词，同时也加载了哈工大停用词词表、四川大学机器智能实验室停用词库、政策特征停用词，以及自建停用词词表进行停用词处理，最后根据词义进行词汇的规范，得到分词处理结果。

（2）关键词提取及变迁分析。本文根据 TF＊IDF 算法对政策文本的分词结果进行关键词提取，并且统计分析了中央科研诚信政策总样本与部分代表部门政策的总体关键词，及其在不同时间段（时间分段依据见3.1小节）的关键词演变情况。

不区分时间段的总体（样本中所有政策文本或某部委机构所有政策文本）关键词分析步骤：首先，用 TF＊IDF 算法筛选出所有文本中的关键词，删去部分意义较小的词（如“加强”“应当”“建设”等）后取 Top15 关键词；然后，计算4个时间段中 Top15 的每个关键词在四个时间切片内的 TF＊IDF 值，用该值来反应关键词在每个时间段中的重要程度；最后，通过折线图形式反映总体的关键词在四个时间段中的变化情况。

区分时间段的关键词分析步骤：首先将样本中的分词结果分时间段存储；然后用 TF＊IDF 算法筛选出各个时间段中关键词，删去部分意义较小的词后取 Top15 关键词；最后对比各时间段关键词的变化情况。

（3）关键词聚类。提取中央科研诚信政策中总 Top50 关键词，去重后使用 scipy 科学计算工具包（https://www.scipy.org）对关键词进行共词分析（按每段文本共现次数计数），使用 VOSviewer 软件对共词分析结果进行聚类，最后将聚类结果交以专家解读添加聚类标签。

2.2　地方政策数据获取及分析方法

由于总数据集中的部分文件例如《科技进步法》涵盖内容较广，仅部分条目中包含科研诚信相关的措施，导致众多响应转发的政策中并非全部都与科研诚信相关，为了使分析的结果更准确紧扣“科研诚信”主题，本研究从108份中央发布的文件中筛选出71份与科研诚信直接相关的文件，用以分析地方政策响应的情况。具体方法如图2所示：

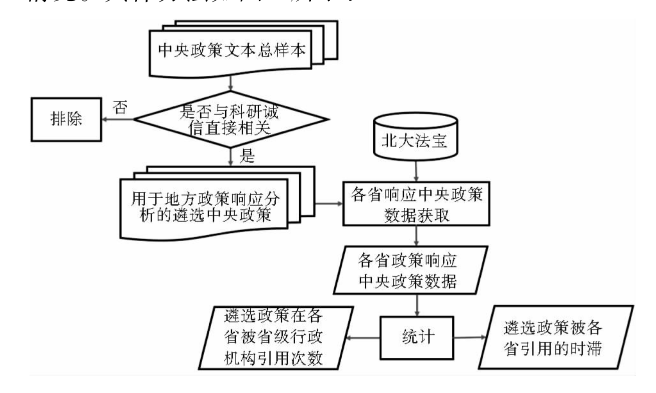


图2　地方政策响应数据及方法

响应数据的获取方面，本文从北大法宝数据库中检索数据集中相应的中央政策，接着获取这些政策在北大法宝数据库中被省级行政机构转发的情况数据，以省为等级，标引记录每个政策被省市机构参照落实、转发的情况。响应时滞 Ti 为 i 省第一次转发政策 j 的年份与政策 j 发布年份之差的平均值。计算公式如下：

$$Ti = \frac{\sum_{j=1}^{n} y_{forward} - y_{publication}}{n} \quad (n \neq 0) \qquad 公式(1)$$

$y_{forward}$：i 省第一次转发参照政策 j 的年份

$y_{publication}$：政策 j 发布年份

n：i 省转发参照科研诚信政策数量

3　结果分析

3.1　科研诚信政策发布时间分析

科研诚信相关政策的年发文量统计结果见图3。从统计结果可以观察到，我国科研诚信政策发文数量整体呈增长的态势。近年来科研不端事件频发，从发

文量的整体增长上可以看到国家对科研不端的治理越来越重视，在有重大科研不端事件时，政策措施也会相应跟进，发文量也会随之增长达到高峰。根据发文量的变化趋势与重大科研不端案例发生时间，我们将科研诚信政策的发布总体分为 4 个时间阶段。1980 – 2001 年，缓慢发展期，政策发文量低，总发文数只有 6 件，有发文的年份数量都在 1 – 2 件之间。2002 – 2006 年，平稳增长期，该阶段发文数达到 20 件，相对上一个阶段，这一时期发文量有明显的增长，每年的发文有 3 – 5 件。接下来的时间有许多重大案例频频曝光，其中有两个影响重大的事件，2006 年上海交通大学陈进"汉芯一号"造假，并借助"汉芯一号"骗取了高达上亿元的科研基金的"汉芯事件"曝光，促成了第二年科技部科研诚信建设办公室的成立，2007 年的发文量直接达到了一个高峰；2011 年，因推荐材料中存在代表著作严重抄袭和经济效益数据不实等问题，科技部通告撤销李连生 2005 年国家科学技术进步奖二等奖获奖项目，第二年相关政策再次骤增，自此总体的科研诚信相关政策发布量总体较密集。这两个重大事件与发文峰值伴随出现，我们因此将 2006 – 2018 年分为两个阶段，2007 – 2011 年，增长波动期，该阶段发文数 26 件，这一阶段的平均发文数量相比上一阶段又有了增长，但每年的数量有所波动；2012 – 2018 年，爆发增长期，这一阶段发文量有了明显的增加，达到 56 件，发文高峰期的年份发文量有 13 件、16 件之多。

总体而言，这四个阶段中的前两个阶段代表的前半段发文较少，数量近似线性地增长，后两个阶段代表的后半段发文骤增，数量近乎指数地增长，也体现了我国对于科研诚信治理的重视程度增加的发展趋势。

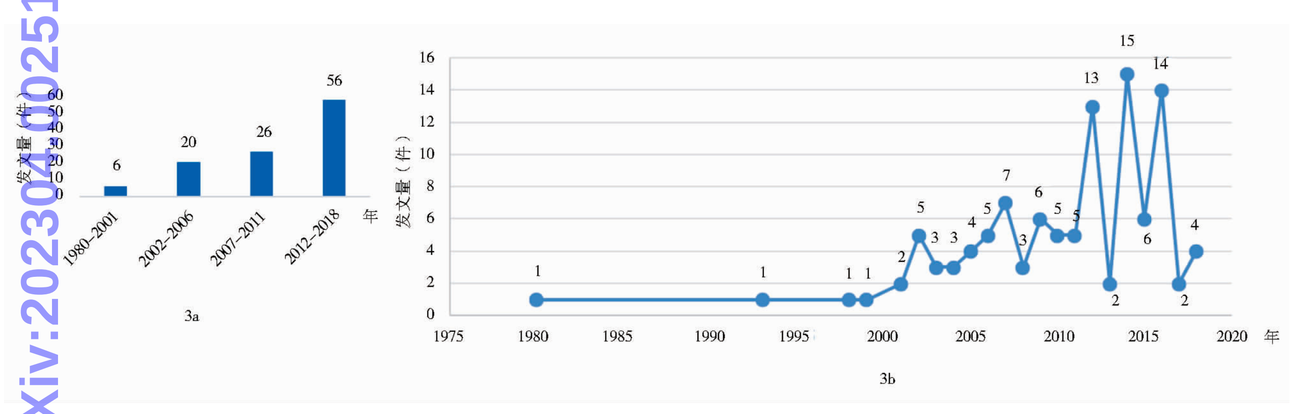


图 3　中央科研诚信相关政策年发文量及阶段分布

3.2　科研诚信政策发布机构及其变迁分析

3.2.1　机构发文及其变迁分析

为了了解各机构发文及其变迁，我们对机构总发文与各阶段发文数量作了统计。机构发文数量由牵头发文（包含独立发文）与联合发文共同组成，结果见图 4，机构发文时序变化见表 1。

机构发文方面，108 件中央政策文件中，一共涉及 21 个机构，其中教育部作为科研诚信治理的主战场发文最多，共发文 35 件，其中牵头发文 26 件，联合发文 9 件；中国科学技术协会作为中央领导下的重要人民团体组织，积极地投入到了科研诚信建设中，一共发文 19 件，其中牵头发文 14 件，联合发文 5 件；中国科学院作为主要的科研投入产出部门对科研诚信也十分重视，一共发文 19 件，其中牵头发文 11 件，联合发文 8 件；科技部作为科研诚信问题的主要治理机构也具有较高的发文数量，一共发文 16 件，其中牵头发文 12 件，联合发文 4 件。在 21 个发文机构中，中共中央与国务院办公厅也有较多的发文，分别发文 12 件与 7 件，表示着国家层面对科研诚信治理的重视。

各机构发文变迁方面，从表 1 可以发现全国人大、中共中央与国务院三个国家权力机构中，全国人大文件发布时间集中在第一阶段，代表着我国相关的法律法规基础较早也有涉及相关的规制，但近年来还没有出台法律法规专门对科研诚信做出约束。中共中央与国务院对科研诚信问题都呈现了越来越关注的趋势，尤其国务院在第四阶段呈现了高度关注，这一阶段集中发文 10 份。

在其他机构中，教育部呈逐渐关注、重视趋势，四个阶段发文越来越多；科技部与中科院在四个阶段中始终关注科研诚信的治理问题，四个阶段发文一直较多；如财政部、人社部等部委虽然不是科研诚信管理的主要机构，但在各阶段也一直配合出台相关政策。

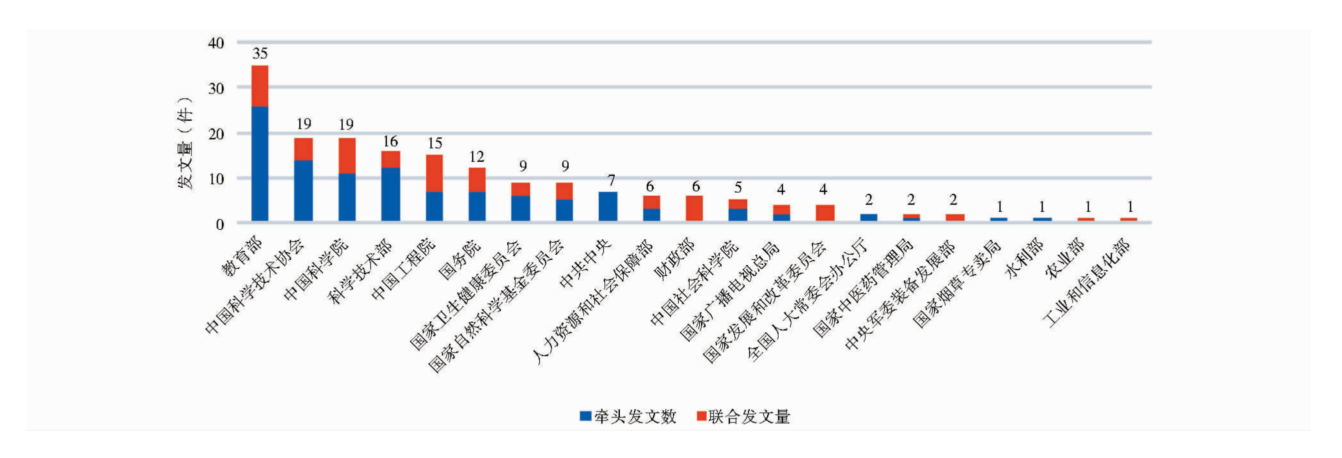


图4　中央科研诚信政策发文数量

注:该图“中共中央”统计项包含发文机构为“中共中央”“中共中央宣传部”“中共中央组织部”“中共中央办公厅”的文件;“国务院”统计项包含发文机构为“国务院”与“国务院办公厅”的文件

表1　中央科研诚信政策发文时序变化

发文机构	1980－2001年	2002－2006年	2007－2011年	2012－2018年	合计
全国人大	2	0	0	0	2
中共中央	0	0	2	5	7
国务院	0	1	1	10	12
教育部	1	7	11	16	35
中国科协	1	0	8	11	20
中科院	3	4	4	8	19
科技部	2	6	5	3	16
工程院	1	1	2	11	15
卫健委	0	0	4	5	9
基金委	0	3	2	4	9
人社部	0	1	2	3	6
财政部	1	2	2	1	6
社科院	0	1	1	3	5
发改委	1	2	0	1	4
广电总局	0	0	0	4	4
军委装备发展部	0	0	1	1	2
中医药管理局	0	0	1	1	2
工信部	0	0	0	1	1
水利部	0	0	1	0	1
国家烟草局	0	0	0	1	1
农业部	0	0	0	1	1

3.2.2　机构发文合作变迁分析

为了了解机构间合作情况的变迁,我们试图统计各阶段的机构合作关系,但由于第一阶段的科研诚信政策发文量较少,分4个阶段统计的合作矩阵前两个阶段结果十分稀疏,因此我们将机构合作关系分为前后两段统计,第一、第二阶段即1980－2006年为前半段,第三、第四阶段即2007－2018年为后半段,分别统计前后两个大阶段的机构合作关系。前半段1980－2006年机构合作网络图见图5,后半段2007－2018年机构合作网络图见图6。网络图中节点大小表示该阶段的机构总发文数量,连边的粗细表示该阶段两个机构间的合作发文频次。

从图5与图6中可以观察到,前后两个阶段的差异非常明显,前半段中各机构间的合作网络还较为稀疏,在后半段明显变得十分稠密。说明在后半段无论是从合作频次还是合作关系上都显著增加,代表了我国各部门越来越趋向于共同治理的现状。

前后两个阶段分开看,前半段中科技部是合作网络的中心,合作数量与合作强度都最突出,到了后半段教育部在合作网络中转为中心地位。在前半段中发文的主要还是与科技活动与科技管理相关的部门,如科技部、教育部、基金委与中科院等,到了后半段更多相关的机构加入进来,主动参与治理。

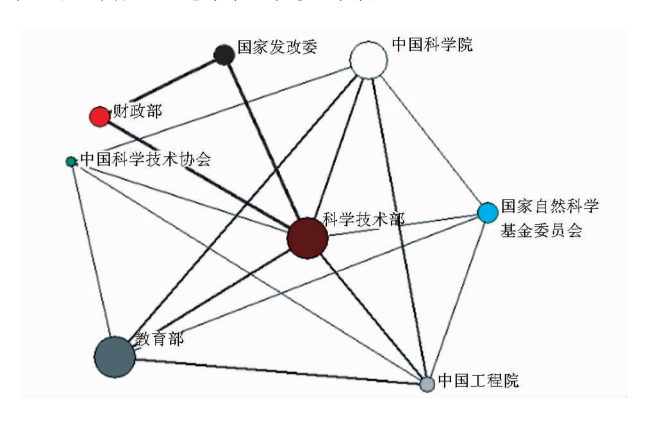


图5　机构发文合作网络1980－2006年

3.3　基于关键词的科研诚信政策内容变迁分析

政府的执政理念与方针内化于政策文献中,通过分析我国科研诚信政策文献内容的变迁,可以了解我国对科研诚信治理关注着力点与方法的变化。在本节中,我们分析了科研诚信政策总体的内容及其变迁,并依据从上文分析中得到的重要的管理部门,即教育部、科技部两个科研管理机构,分析了其政策关注点及其变化,更加详细地描绘了我国科研诚信治理的图景。

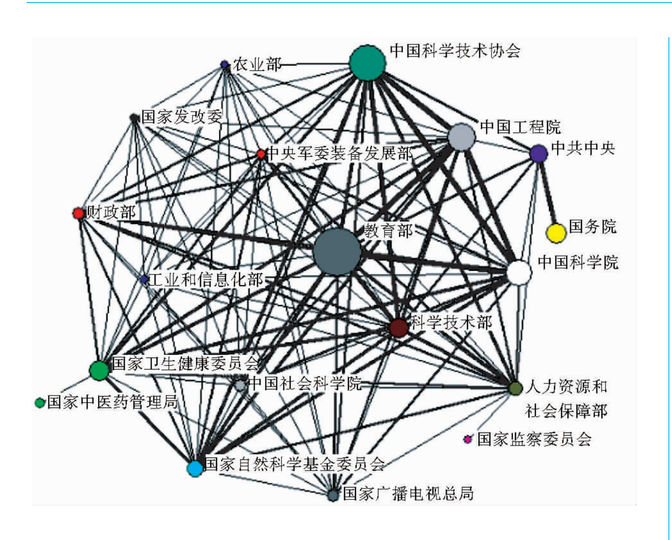


图 6 机构发文合作网络 2007 – 2018 年

3.3.1 中央科研诚信政策内容分析

为了发现我国的科研诚信政策主要内容，我们使用 2.1.3 小节中的方法，对中央科研诚信政策中的关键词做了聚类分析，聚类结果如图 7 所示，根据分析与解读，我们发现我国科研诚信政策内容主要关注学风建设、项目管理、信用监督以及人才创新这四个方面，这些主题表现了我国科研诚信政策的工作建设重点，从源头上我们要进行“学风建设”，从成果上我们需要进行“项目管理”，改善评价评审等方式，从科研人员方面要鼓励“人才创新”，同时也要建立完善的“信用监督”等制度。

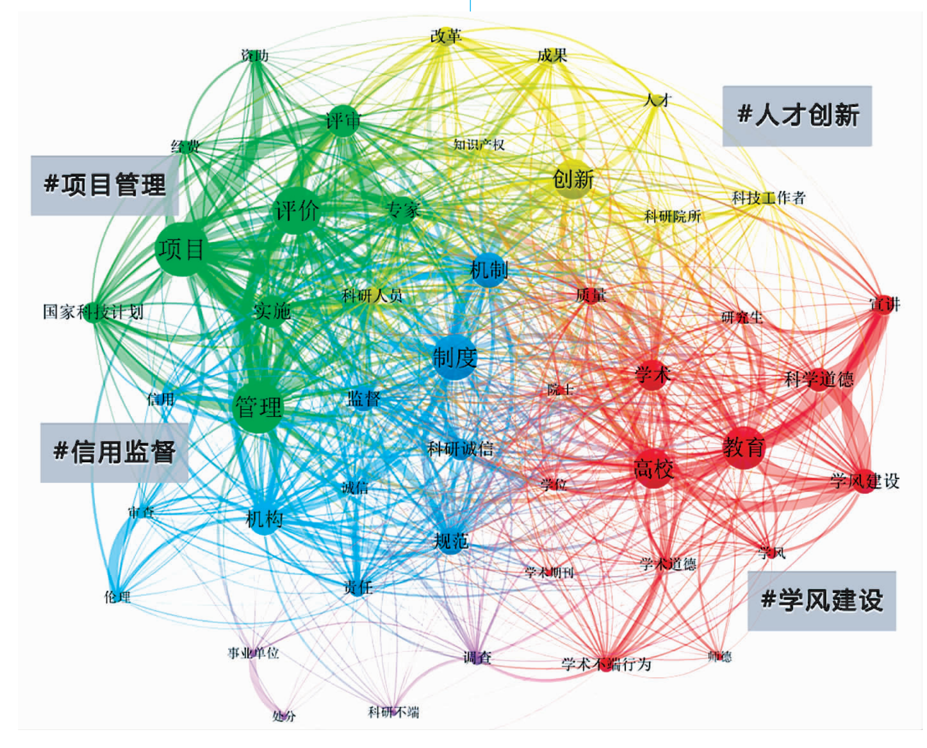


图 7 中央科研诚信政策关键词聚类

为了解总体而言有哪些重要的内容单元，我们统计了我国中央科研诚信政策文本中重要程度 Top15 的关键词（见表 2），这些重要关键词在四个时间段中的变化情况见图 8。

总体而言，我国科研诚信政策围绕着“科研”“科技”与“项目”展开，“科研”与“科技”二词在四个阶段中都较重要，而“项目”一词的重要性相对而言逐渐降低，或可说明我国对于科研诚信的治理对象有所转移；科研诚信的治理与“评审”“评价”机制密切相关，这两个词具有相同的趋势变化，我国的科研诚信政策在第二阶段 2002 – 2006 年比较明显地重视科研评价制度，出台了《关于改进科学技术评价工作的决定》为代表的一系列规定；“科研诚信”一词虽然在第一阶段 1980 – 2001 年还未进入政策的话语体系，但自第二阶段加入以后重要性连年上升；“创新”一词重要性的上升势头也十分明显，科研诚信是科技创新的基石[21]，为了创新需要进行科研诚信治理，进行科研诚信治理可以促进创新；在治理科研诚信的方式上，“学风建设”“教育”与“宣讲”等词语都具有一定重要性，“学风建设”在第三阶段 2007 – 2011 年重要程度达到峰值，“教育”与“宣讲”的重要性也一直在上升中，说明进行学风建设与宣讲教育都作为了我国尝试科研诚信治理的方法手段。

冯凌子，刘敬，袁军鹏. 我国科研诚信政策变迁计量分析[J]. 图书情报工作,2020,64(9):73－84.

表 2　Top15 中央科研诚信政策关键词

序号	关键词	TF＊IDF
1	科研	0.117 3
2	项目	0.073 5
3	科技	0.065 5
4	评价	0.062 5
5	评审	0.056 7
6	创新	0.055 1
7	高校	0.047 9
8	学风建设	0.043 9
9	科研诚信	0.042 4
10	教育	0.041 5
11	国家科技计划	0.040 2
12	宣讲	0.037 8
13	科学道德	0.036 8
14	学术	0.036 7
15	管理	0.033 5

3.3.2　中央科研诚信政策内容变迁分析

为进一步了解各阶段内容变化的情况，我们统计了四个阶段 Top15 的关键词（如表 3 所示），并将词的变化情况用色块标注。其种每列灰色的色块代表下阶段将消失的词汇，橙色的色块代表本阶段新增的词汇，蓝色的色块代表本阶段新增且下阶段消失的词汇。

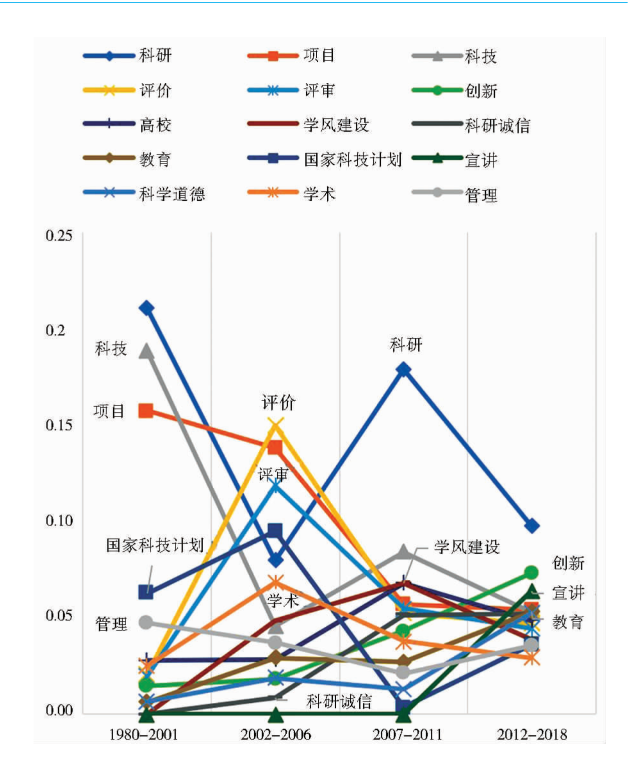


图 8　Top15 中央科研诚信政策关键词重要程度变化

表 3　中央科研诚信政策各时间阶段关键词 Top15

序号	1980－2001 年		2002－2006 年		2007－2011 年		2012－2018 年	
	关键词	TF＊IDF	关键词	TF＊IDF	关键词	TF＊IDF	关键词	TF＊IDF
1	科研	0.210 5	评价	0.149 9	科研	0.178 8	科研	0.097 5
2	科技	0.188 2	项目	0.138 0	科技	0.084 1	创新	0.073 0
3	项目	0.157 5	评审	0.118 2	高校	0.067 8	宣讲	0.063 8
4	学位	0.069 2	国家科技计划	0.095 0	学风建设	0.067 7	项目	0.053 9
5	科技人员	0.064 7	科研	0.079 6	伦理	0.062 5	教育	0.053 1
6	科技工作者	0.062 5	学术	0.068 2	科学	0.060 1	科技	0.052 8
7	国家科技计划	0.062 5	学术道德	0.065 4	科技工作者	0.058 7	科学道德	0.052 7
8	财政性	0.054 8	学风建设	0.048 3	项目	0.057 1	科研诚信	0.052 3
9	学部	0.053 4	科技	0.045 6	评审	0.054 8	信用	0.048 7
10	经费	0.050 8	师德	0.044 5	评价	0.052 2	高校	0.048 3
11	管理	0.047 2	招标	0.044 4	科研诚信	0.051 4	评价	0.046 9
12	科技进步	0.040 7	投标	0.037 6	资助	0.051 3	评审	0.044 5
13	预算	0.040 3	通报批评	0.037 4	委员会	0.050 6	学风建设	0.038 9
14	投诉	0.037 9	管理	0.0371	学术不端行为	0.043 5	国家科技计划	0.036 0
15	技术	0.037 4	招聘	0.032 2	创新	0.043 0	管理	0.035 6

从各阶段内容变化上看，我们直观地看到随着时间变化，每个阶段的政策重点内容都在改变，直到第四阶段新增的重要词汇才逐渐相对稳定下来，第一阶段与第四阶段的关键词已经完全不同。各个阶段变化大，表征着治理的热点难题在变化。

从各阶段内容热点上看，每个阶段都有随着相应事件催生的对应政策。第一阶段 1980－2001 年发布了《中华人民共和国学位条例》《中华人民共和国科学技术进步法》等法律法规，《关于科技工作者行为准则的若干意见》等规范文件主要将重点放在科学工作者的自我约束上。第二阶段 2002－2006 年发布了《关于加强学术道德建设的若干意见》《关于改进科学技术评价工作的决定》等文件，从学术道德建设与科技评价方式改进等方面治理科研诚信问题，这一时间段我国还发布了《关于国家科研计划实施课题制管理的规定》实施了科研计划管理的改革，对科研诚信的管理规

制内容更加地具体了。第三阶段 2007－2011 年发布了《关于加强我国科研诚信建设的意见》《关于严肃处理高等学校学术不端行为的通知》《切实加强和改进高等学校学风建设的实施意见》等文件,对科研不端行为有了界定,更加明确了对科研诚信的治理任务,方法主要是从学风建设与改进项目评审等方面进行治理。第四阶段 2012－2018 年重要文件频发,《学位论文作假行为处理办法》《高等学校预防与处理学术不端行为办法》《国家科技计划(专项基金等)严重失信行为记录暂行规定》《关于进一步加强科研诚信建设的若干意见》等文件给出了更加具体的科研诚信建设目标与惩处规制的具体处理办法依据。

3.3.3　重要机构科研诚信政策内容分析

两个重要机构的重要内容如表 4 所示,从中可以发现,教育部与科技部的科研诚信政策关注点与着力点有着明显的职能特色,教育部重"学风建设",重"师德"建设,对科研诚信的治理更偏向"师－生－学术"这一生态,对不端的定义更偏向"学术不端";科技部管理科研活动,主体涉及"国家科技计划"等"项目",以及"科技工作者",对科研诚信的治理更偏向"国家－科研人员－项目"这一生态,对不端的定义更偏向"科研不端"。

表 4　重要机构科研诚信政策关键词 Top15

序号	教育部		科技部	
	关键词	TF * IDF	关键词	TF * IDF
1	高校	0.132 4	项目	0.200 6
2	科研	0.119 8	评价	0.161 8
3	评价	0.108 4	国家科技计划	0.157 3
4	学风建设	0.096 4	评审	0.119 8
5	科技	0.073 8	科技	0.109 5
6	教育	0.071 5	科技工作者	0.095 2
7	学术	0.068 7	科研	0.080 2
8	师德	0.062 1	科研诚信	0.069 0
9	学术不端行为	0.060 0	管理	0.059 8
10	科技工作者	0.053 1	招标	0.049 8
11	学术道德	0.053 1	科研不端	0.044 6
12	学位	0.049 6	投标	0.042 1
13	创新	0.042 5	创新	0.038 2
14	社会科学	0.041 8	招标人	0.034 9
15	宣讲	0.039 7	经费	0.034 9

这些关键词的重要性变化情况如图 9、图 10 所示。两个机构由于在第一阶段发文少,这一阶段词都较少,因此第二阶段都新增了许多关键词。总观四个阶段的情况,教育部"学术不端行为"一词在多年政策演进中重要性越来越高,对于"学术道德""师德"的建设在第二阶段最为明显,之后有所下降,"学风建设"自第二阶段被提出后一直较为重要;科技部"科研诚信"一词自第二阶段进入话语体系后一直较重要,在第二阶段达到峰值,第二阶段的重要词"评价"与"评审"、第四阶段的"国家科技计划"与"经费"等词都体现着各个阶段的工作对象重点。

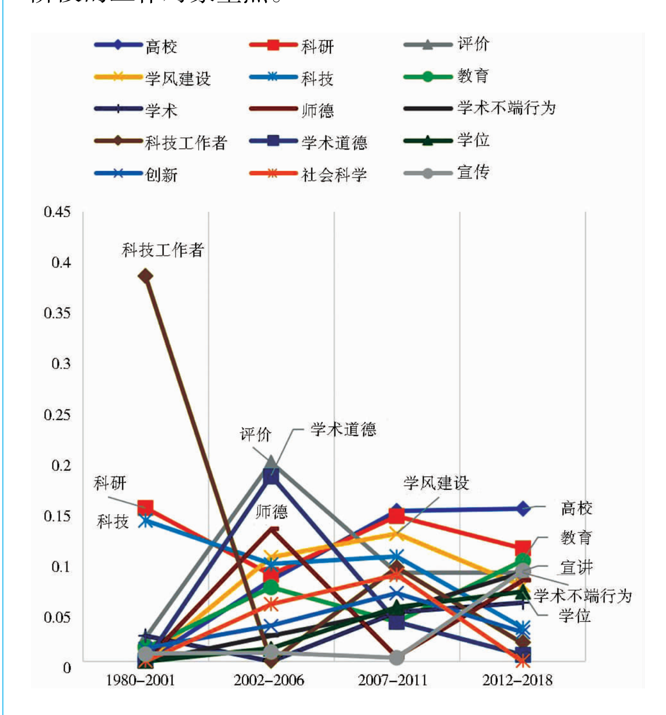


图 9　Top15 教育部政策关键词重要程度变化

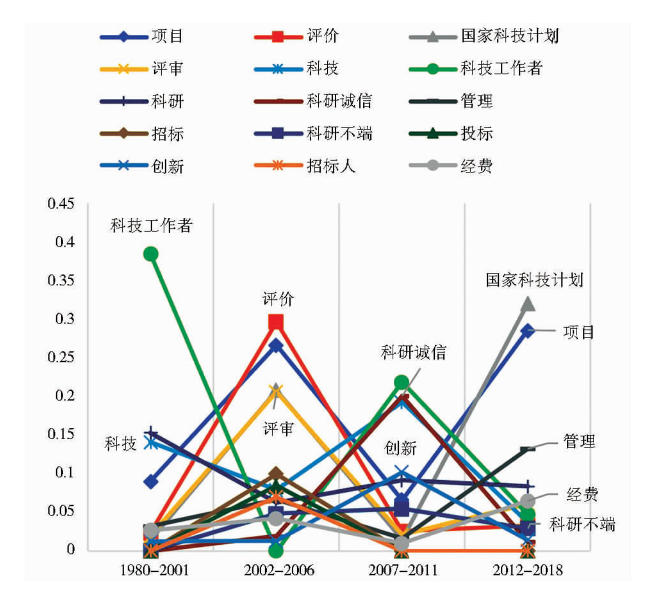


图 10　Top15 科技部政策关键词重要程度变化

3.3.4　重要机构科研诚信政策内容变迁分析

表 5 与表 6 是两个科研诚信治理重要机构教育部与科技部在各个时间段的重要关键词,两个部门从该

冯凌子，刘敬，袁军鹏. 我国科研诚信政策变迁计量分析[J]. 图书情报工作,2020,64(9):73－84.

表中我们能够更加直观地看到各个时段政策内容的变化,两个重要机构之间工作的差异与各自的变化都十分明显,但与上文提到的中央科研诚信政策总样本中的变化趋势是一致的。值得说明的是,两个机构由于在第一阶段1980－2001年仅有一份二者都参与其中的文件《关于科技工作者行为准则的若干意见》,因此在第一阶段两个机构的结果是完全一样的。

表5　教育部各时间阶段关键词Top15

序号	1980－2001年		2002－2006年		2007－2011年		2012－2018年	
	关键词	TF * IDF	关键词	TF * IDF	关键词	TF * IDF	关键词	TF * IDF
1	科技工作者	0.385 2	评价	0.198 7	高校	0.150 2	高校	0.152 4
2	科研	0.153 3	学术道德	0.185 3	科研	0.145 0	科研	0.112 6
3	科技	0.140 9	师德	0.132 2	学风建设	0.127 5	教育	0.100 6
4	项目	0.089 9	学术	0.132 0	科技	0.104 6	宣讲	0.090 4
5	科学	0.089 9	学风建设	0.104 2	科技工作者	0.093 6	学术不端行为	0.088 6
6	伪科学	0.064 0	科技	0.097 2	评价	0.088 9	评价	0.088 5
7	科技成果	0.058 5	科研	0.086 5	社会科学	0.086 6	师德	0.080 6
8	迷信活动	0.053 9	高校	0.081 3	科研诚信	0.080 1	学风建设	0.075 6
9	不良行为	0.05 18	教育	0.073 7	创新	0.067 8	研究生	0.074 4
10	技术	0.048 9	教师	0.072 5	科技委	0.066 8	学位	0.069 4
11	社会	0.046 5	社会科学	0.057 2	学部	0.056 0	学位论文	0.062 7
12	成果	0.046 3	论文	0.051 1	哲学	0.055 9	学术	0.058 3
13	自觉	0.042 6	学术规范	0.048 6	学位	0.054 0	项目	0.055 2
14	科技知识	0.040 8	制度	0.045 0	学术不端行为	0.050 2	科技期刊	0.055 1
15	管理机构	0.040 7	学术风气	0.044 0	学术	0.049 1	科研人员	0.038 5

从教育部的重要关键词变化中,我们可以看到部门的政策重点具有一定的继承性,后三个阶段围绕着“学风建设”,从“学术道德”与“学术规范”,到开始注重“创新”与发现“学术不端行为”,再到进行“宣讲”,关注“研究生”“学位论文”与“科研人员”“科技期刊”论文,都体现了教育部科研诚信治理工作中的逐步探索与重点变化。

科技部的关键词变化相较教育部更加明显,其在第二阶段就较早地关注到了“科研不端”,有着从“评审”、“评价”与“招投标”,到关注“创新”、“宣传”与科研人员,再到对“失信”“责任”界定的变化。特别是第4阶段不再强调“科研不端”,改为“失信”。

表6　科技部各时间阶段关键词Top15

序号	1980－2001年		2002－2006年		2007－2011年		2012－2018年	
	关键词	TF * IDF	关键词	TF * IDF	关键词	TF * IDF	关键词	TF * IDF
1	科技工作者	0.385 2	评价	0.296 3	科技工作者	0.218 9	国家科技计划	0.321 1
2	科研	0.153 3	项目	0.266 5	科研诚信	0.200 7	项目	0.286 1
3	科技	0.140 9	国家科技计划	0.208 2	科技	0.192 8	管理	0.132 4
4	项目	0.089 9	评审	0.206 6	创新	0.102 2	失信	0.126 3
5	科学	0.089 9	招标	0.099 7	科研	0.091 7	论文	0.103 7
6	伪科学	0.064 0	投标	0.084 4	宣传	0.083 1	法人	0.093 1
7	科技成果	0.058 5	科技	0.079 4	项目	0.065 8	责任	0.091 6
8	迷信活动	0.053 9	招标人	0.070 0	科学	0.063 8	项目管理	0.084 6
9	不良行为	0.051 8	管理	0.066 3	创新型	0.060 9	科研	0.084 2
10	技术	0.048 9	科研	0.063 1	科研不端	0.054 7	记录	0.076 0
11	社会	0.046 5	投标人	0.056 9	科技事业	0.044 2	经费	0.064 9
12	成果	0.046 3	专家	0.050 8	专项	0.042 8	评审	0.061 5
13	自觉	0.042 6	评标	0.050 7	科技人员	0.041 8	五不准	0.052 5
14	科技知识	0.040 8	科研不端	0.049 0	自主	0.041 3	违反	0.050 6
15	管理机构	0.040 7	经费	0.042 4	制度	0.036 8	科技工作者	0.046 7

3.4 科研诚信政策地方响应分析

在中央部委等国家机构出台政策对科研诚信问题实施治理后，需要地方的各个相关机构相应落实，从而达到治理的效果。因此可以从全国省市政策参照、转发的情况来大致观测国家层面的政策被落实的情况。

我们统计了各省对于中央政策（经筛选后与科研诚信直接相关的政策）的响应情况与响应时滞，响应“数量 – 时滞”分布如图 11 所示。从响应时滞上看，各省响应时间从当年到十年不等，平均响应时长为 2.65 年。政策响应“数量 – 时滞”分布的散点图中右上角的点为响应数量多、响应速度快的省市，左上角为响应较少、响应速度快的省市，左下角为响应较少、响应速度慢的省市。我们可以观察到图中右上角区域上海、北京、浙江、湖南这些省响应多、速度快，吉林、河北等省市响应快、数量适中。

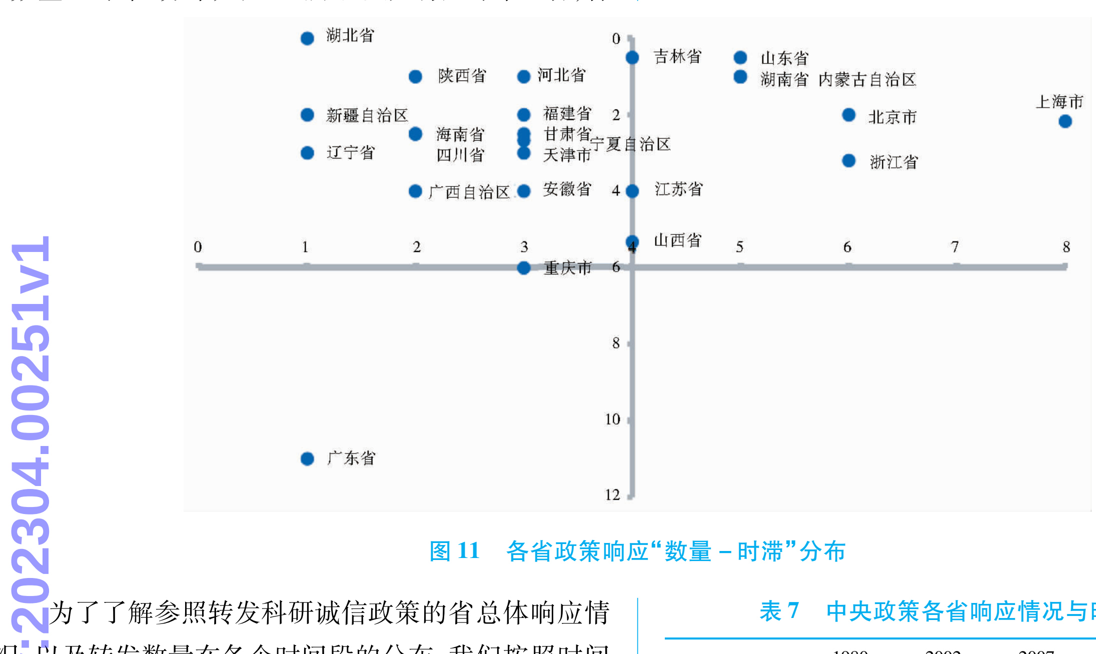


图 11　各省政策响应“数量 – 时滞”分布

为了了解参照转发科研诚信政策的省总体响应情况，以及转发数量在各个时间段的分布，我们按照时间段统计了各省发文的数量并进行了数量汇总，结果如表 7 所示。从响应总数上看，上海的参照转发落实次数最多，一共有 8 份文件，北京与浙江紧随其后，各有 6 份文件转发，湖南、内蒙与山东的转发也较多，各有 5 份文件转发；另外我国各省对于科研诚信政策的转发落实情况覆盖面很大，从目前被统计到的数据来看，全国有 75% 的省都转发了中央的科研诚信政策，积极参与到科研诚信的治理中来。从各阶段响应情况上看，从表中统计结果我们发现，响应主要集中在后两个阶段，其中大部分省的响应都集中在第四阶段，且几乎所有省在第四阶段都有发文；有一半的省在第三阶段有发文采取措施进行科研诚信治理；仅有两个省在第二阶段开始就有发文，从目前统计到的情况来看，还没有省在初期第一阶段转发。这一总体数量趋势与国家层面的发文数量与文件重视程度相一致，但是相对而言略微滞后。

由于本文对政策地方响应情况的统计数据仅从北大法宝数据库收录的省市级机构政策参照、转发情况作了统计，可能有一些参照转发未被收录，例如江西省

表 7　中央政策各省响应情况与时间分布

发文省	1980 – 2001 年	2002 – 2006 年	2007 – 2011 年	2012 – 2018 年	合计
上海市	0	0	1	7	8
北京市	0	0	2	4	6
浙江省	0	0	2	4	6
湖南省	0	0	0	5	5
内蒙古自治区	0	0	1	4	5
山东省	0	0	3	2	5
吉林省	0	0	1	3	4
江苏省	0	0	1	3	4
山西省	0	1	0	3	4
甘肃省	0	0	0	3	3
天津市	0	0	0	3	3
安徽省	0	0	1	2	3
福建省	0	0	1	2	3
河北省	0	0	1	2	3
重庆市	0	1	0	2	3
宁夏自治区	0	0	2	1	3
海南省	0	0	0	2	2
陕西省	0	0	0	2	2
四川省	0	0	0	2	2
广西自治区	0	0	1	1	2
广东省	0	0	0	1	1
辽宁省	0	0	0	1	1
新疆自治区	0	0	0	1	1
湖北省	0	0	1	0	1

教育厅发布的《江西省高等学校学术不端行为调查处理规程》;各个省份还可能发布了一些未标明具体参照转发的文件,也贯彻落实了科研诚信治理的精神,这些实践暂时还没有在本文中得到体现,例如河南省科技厅发布了《河南省省级科技计划不良信用行为记录暂行规定》等。因此结果仅作为参考,只能从一定程度上的反映响应的现状。

4 结论与讨论

本文搜集整理了 108 份中央发布的与科研诚信相关的政策,及 80 份参照落实中央政策的地方政策,使用政策文本计算方法,分析了我国科研诚信政策总体以及变迁的情况。通过对政策发布时间、发布机构、政策内容及政策响应情况的现状及其变迁进行的具体分析梳理,得到了政策发布的时间特征、重要机构特征、内容侧重点变化以及政策响应的特征。我们发现:①科研诚信相关治理措施早在 1980 年就有涉及,但处理措施办法还不太具体。近年来围绕国家创新建设的需求发文数量骤增,治理的态度更坚决,措施办法也更加具体。虽然及至 2018 年中才出台与科研诚信直接相关的法规,但目前也还没有科研诚信立法,这可能会造成对于具体案例的处理上标准不一的情况。②科研诚信的治理由中央政府领导,教育部与科技部是各个阶段的主要管理部门,相关的一些部门随着时间推移也逐渐加入进行共同治理,目前已形成合力的格局。③各部门的科研诚信治理着力点、对象与方法略有差异,但总体而言方法与目标上逐渐从自我约束、道德教育转变到制度建设与道德建设齐抓。④全国范围内绝大多数的地方省都积极响应了中央政策,采取措施进行科研诚信治理,但相对而言有一定的时滞。

与已有的相关研究[17－18]相比,方法上,本文不仅使用了政策计量方法,还结合了文本计算方法,对我国科研诚信政策内容层面的内容与变迁进行了揭示;分析内容上,本文不仅分析了科研诚信政策发文机构,还分析了这些机构的合作状况与重要机构的政策内容;以及它们的变迁,除此之外,本文还分析了地方省对于国家政策的响应情况,有助于了解我国科研诚信政策颁布后响应及落实的概况。

但研究目前还存在以下不足:①政策文本数据来源待丰富,部分地方响应政策数据的缺失,或会影响数据结论;②由于地方响应数据量较小,目前地方政策未区分政策转发(扩散)、政策参照落实与隐性响应,反映的结果存在与现状有部分偏差的可能;③对结果的解读目前还停留在数据层面。针对这些不足,我们下一步工作中会采取以下改进方式:扩充数据来源,使用爬虫等方式抓取更全的相关政策文本;区分地方政策响应政策类型,深入发掘信息;进行专家访谈,对数据结果进行深入解读。

参考文献:

[1] 结果难服众 学界期待公布韩春雨事件调查细节[EB/OL].[2019－04－30]. http://news.sciencenet.cn/htmlnews/2018/9/417149.shtm.

[2] STIGBRAND T. Retraction Note to multiple articles in Tumor Biology[EB/OL].[2019－04－30]. https://doi.org/10.1007/s13277-017-5487-6.

[3] 南大公布梁莹处理意见:取消研究生导师资格[EB/OL].[2019－04－30]. http://js.people.com.cn/n2/2018/1215/c360307-32414128.html.

[4] 北电成立"翟天临事件调查组"[EB/OL].[2019－04－30]. http://epaper.ynet.com/html/2019-02/12/content_319264.htm?div=-1.

[5] 刘延东:加强科研诚信学风建设 促科学事业健康发展[EB/OL].[2019－04－30]. http://www.gov.cn/ldhd/2010-03/30/content_1569423.htm.

[6] 习近平:决胜全面建成小康社会 夺取新时代中国特色社会主义伟大胜利——在中国共产党第十九次全国代表大会上的报告[R]. 北京:人民出版社, 2017:1－71.

[7] 2019 年政府工作报告全文[EB/OL].[2019－04－30]. http://www.gov.cn/zhuanti/2019qglh/2019lhzfgzbg/index.htm.

[8] 黄萃. 政策文献量化研究[M]. 北京:科学出版社, 2016.

[9] 裴雷,孙建军,周兆韬. 政策文本计算:一种新的政策文本解读方式[J]. 图书与情报, 2016(6):47－55.

[10] HUANG C, SU J, XIE X, et al. A bibliometric study of China's science and technology policies:1949－2010[J]. Scientometrics, 2014, 102(2):1521－1539.

[11] 张剑,黄萃,叶选挺,等. 中国公共政策扩散的文献量化研究——以科技成果转化政策为例[J]. 中国软科学, 2016(2):145－155.

[12] 李燕萍,吴绍棠,郜斐,等. 改革开放以来我国科研经费管理政策的变迁、评介与走向——基于政策文本的内容分析[J]. 科学学研究, 2009, 27(10):1441－1447,1453.

[13] 杨凯瑞,何忍星,钟书华. 中国中央政府支持创新创业发展政策文本量化研究(2003－2017 年)——来自国务院及 16 部委的数据分析[J]. 科技进步与对策, 2019, 36(15):107－114.

[14] ZHANG X, YANG F. Rural informatization policy evolution in China:a bibliometric study[J]. Scientometrics, 2019, 120(1):129－153.

[15] 黄萃,苏竣,施丽萍,等. 政策工具视角的中国风能政策文本量化研究[J]. 科学学研究, 2011, 29(6):876－882,889.

[16] 袁军鹏,淮孟姣,潘云涛,等. 我国科研诚信研究发展概述:科学计量学视角[J]. 国防科技, 2017, 38(6):14－20.

[17] 田瑞强，姚长青，刘泙颖，等．学术不端治理政策及案例计量研究[J]．中国科技期刊研究，2018，29(4)：355－361．
[18] 靳彤，张红伟，赵勇．政策文本计量视角下我国科研诚信治理的特征与启示[J]．情报工程，2018，4(5)：116－126．
[19] 黄萃，任弢，张剑．政策文献量化研究:公共政策研究的新方向[J]．公共管理学报，2015，12(2)：129－137，158－159．
[20] 科学技术部科研诚信建设办公室．科研诚信建设相关法律法规和文件汇编[M]．北京：高等教育出版社，2017．
[21] 中共中央办公厅 国务院办公厅印发《关于进一步加强科研诚信建设的若干意见》[EB/OL]．[2019－04－30]．http://www.gov.cn/zhengce/2018-05/30/content_5294886.htm．

作者贡献说明：

冯凌子:设计论文框架,数据处理与分析,撰写论文;
刘敬:政策文本数据处理;
袁军鹏:提出论文思路,参与研究并修订论文。

Quantitative Analysis of the Changes of Research Integrity Policy in China

Feng Lingzi[1,2]　Liu Jing[3,4]　Yuan Junpeng[1,2]
[1] National Science Library, Chinese Academy of Sciences, Beijing 100190
[2] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190
[3] Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093
[4] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049

Abstract: [**Purpose/significance**] This paper analyses the current situation and changes of the research integrity policy issued by the Chinese central government and the local response, in order to grasp the development trend and stage focus of the policy, provide support for the construction of research integrity, and provide help for the development of research integrity policy in China. [**Method/process**] This paper collected 108 policies related to research integrity issued by the central government and ministries, and 80 local policies for the implementation of the central policy. By using the policy text computing and policy documents quantitative research method, this paper analyzed and combed the current situation and changes of policy release time, publishing institutions, policy content and policy response, and obtained the time characteristics of policy release. Finally, this paper got the characteristics of the time of policy release, the characteristics of important institutions, the change of content focus and the characteristics of policy response. [**Result/conclusion**] ①Relevant governance measures of research integrity have been involved earlier. In recent years, China has put forward the need to build an innovative country, so the number of policy papers has increased sharply, the attitude to the governance of research integrity is more resolute, and the measures are more specific. ②The management of scientific research integrity is led by the central government. The Ministry of Education and the Ministry of Science and Technology are the main administrative departments at all stages. Relevant departments have gradually joined in the joint management. At present, a pattern of joint efforts has been formed. ③There are slight differences in the focus, object and method of research integrity governance among different departments, but generally speaking, the trend is to gradually shift from self-restraint, moral education to system construction and moral construction. ④The vast majority of local provinces and municipalities throughout the country have responded positively to the central policy and taken measures to improve scientific research integrity, but there is a relative time lag.

Keywords: research integrity　policy text computing　policy documents quantitative research　policy change　policy analysis